



Project no. 018340

Project acronym: EDIT

**Project title: Toward the European Distributed Institute of
Taxonomy**

Instrument: Network of Excellence

Thematic Priority: Sub-Priority 1.1.6.3: "Global Change and Ecosystems"

C5.127 Prototype of service to generate printed publication from the CDM

Due date of component: Month 49
Actual submission date: Month 50

Start date of project: 01/03/2006

Duration: 5 years

Organisation name of lead contractor for this component: 9 FUB-BGBM

Revision final

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

The prototype service to generate printed publications from the CDM (Printservice) consists of an interface that lets a user select a classification or specific parts of a classification, a system to harvest the required data from the CDM REST Services (Webservice) and XSL stylesheets to process the acquired data into the Open Document Format (ODF). Being a prototype, the system does not claim to be complete, correct or useful in production use.

Prerequisites

The Printservice prototype user interface was realized as eclipse plugin and therefore must be run in an eclipse RCP context, like the EDIT Taxonomic Editor. The harvesting system relies on an accessible CDM Server instance.

A resource to store the generated output has to be accessible by the executing user.

Output Formats

The generation of documents for specific output formats is highly configurable. Implementations that generate different output formats (Output Module) may be plugged into the system in a modular fashion. An Output Module consists of algorithms and instructions to process the raw XML data harvested from the Webservice and mould it into a structure that conforms to a specific output format.

The Printservice prototype includes implementations for two distinct output formats:

- XML: Unaltered output from the Webservice to allow for further processing in third party XSL applications
- ODF: Output from the Webservice processed by XSL and transformed into an ODF document. Certain parts of the document are marked with custom styles exemplarily to showcase the ability to manipulate the layout of the document individually by adjusting ODF text styles.

Workflow

Running the Printservice consists of two separate operations that may be performed independently.

Creating a configuration

A Printservice configuration is a data structure that holds information about:

- A list of taxa that should be processed. This will include all taxonomically included taxa
- A list of Output Modules
- Resource to save the generated output to

It is not specified how the configuration data structure will be created. Possible ways may include parsing of XML configuration files or java property files, form based user interfaces and so on.

The Printservice prototype however, offers a form based user interface to create such a configuration conveniently.

Executing the Printservice

Based on a given configuration, the Printservice performs the stages of harvesting and processing of data.

Harvesting

Based on the list of taxa defined in the configuration, the Webservice is queried to generate a XML DOM that holds the required data.

Processing

The generated XML DOM is passed to the Output Module(s) defined by the configuration. Each Output Module processes the data and stores the resulting document to the resource defined in the configuration.

Results and Conclusion

The Printservice prototype in its current form demonstrates the feasibility of the chosen approach. Tapping the Webservices for retrieving the needed input and processing that input performs well and in a timely manner.

As proposed by the functional description of the prototpye [1], the next steps will be to aim at completeness of data and refinement of the stylesheet that defines the resulting ODF document.

[1] <http://dev.e-taxonomy.eu/trac/wiki/PrintPublisher>